

# ***U.S. PATENT APPLICATION***

***Inventor(s):*** Christiane OSTERHOFF  
Richard IVELL

***Invention:*** EPIDIDYMIS-SPECIFIC RECEPTOR PROTEIN

***NIXON & VANDERHYE P.C.***  
ATTORNEYS AT LAW  
1100 NORTH GLEBE ROAD  
8<sup>TH</sup> FLOOR  
ARLINGTON, VIRGINIA 22201-4714  
(703) 816-4000  
*Facsimile* (703) 816-4100

## ***SPECIFICATION***

Epididymis-specific receptor protein and its use

The invention relates to a novel transmembrane receptor protein which is specific to the mammalian epididymis, DNA sequences which code for the 5 protein and specific ligands, and to the use thereof for the preparation of agents for diagnosis of male infertility and for controlling maturation of spermatozoa.

BACKGROUND OF THE INVENTION

10

About 15% of all couples in the Federal Republic of Germany are currently involuntarily childless, and their proportion is increasing constantly. The reasons for the infertility lie equally with the men and women. However, while the reasons in women have been extensively investigated and can be diagnosed, in men organic causes can be established in only approx. 30% of cases. About one third of the remaining 70% of cases can be attributed to oligospermia of unknown origin, while the causes in the remaining cases are still unclear according to the current state of knowledge. In cases of idiopathic infertility in particular, in which no fertilization of the female 15 ovocyte takes place in spite of a sufficient number of spermatozoa in the ejaculate (H.W.G. Baker et al. (1986) Relative incidence of etiological disorders in male infertility. In: *Male Reproductive Dysfunction* (ed. R.J. Santen and R.S. Swerdloff), p. 341 - 372, Marcel Dekker Inc., New York), a disturbance in the maturation of spermatozoa which takes place in the 20 epididymis is suspected.

25

The epididymis plays a key role in maturation of spermatozoa. In man, it consists of a single tubule about 5 m long, wound up in a meandering form. The still immature spermatozoa formed in the testis are transported to the 30 epididymis, and undergo a maturation process during a passage lasting 2 to 6 days (Moore et al., Fertilizing capacity of rat spermatozoa is correlated

with decline in straight-line velocity measured by continuous computer-aided sperm analysis: epididymal rat spermatozoa from the proximal cauda have a greater fertilizing capacity in vitro than those from the distal cauda or vas deferens, *Journal of Andrology*, 17, 50-60 (1996)). The special medium of the 5 epididymis ensures that the spermatozoa embedded therein remain vital for a long time, and furthermore acquire the particular physical, immunological and biochemical properties of spermatozoa which are capable of fertilization.

It is accordingly to be assumed that in numerous cases male infertility, and 10 in particular idiopathic infertility, is based on disturbances in the maturation processes which take place in the epididymis.

Although the anatomical fine structure of the human epididymis has been described very thoroughly macroscopically, microscopically and by electron microscopy (cf. A.F. Holstein: *Morphologische Studien am Nebenhoden des Menschen, Zwanglose Abhandlungen aus dem Gebiet der normalen und pathologischen Anatomie* [Morphological studies on the epididymis of man, informal papers from the field of normal and pathological anatomy], 20, 1-91 15 (1969)), its protein products have been researched only little, with few exceptions (cf., for example, Tezon et al., *Immunochemical localization of 20 secretory antigens in the human epididymis and their association with spermatozoa*, *Biology of Reproduction*, 32, 591-597 (1985)). Almost all the findings on this organ originate from work on rats, mice, hamsters, boars, bulls or occasionally monkeys, and animal-specific differences are known to 25 be great.

The findings on maturation of spermatozoa during passage through the epididymis obtained on various species of animals can be summarized as follows (cf. T.G. Cooper: *The Epididymis, Sperm Maturation and Fertilization*, 30 Springer Verlag, Berlin, (1986)):

(a) Development of an orientated forward mobility and capability for hyperactivation of the spermatozoa.

5 (b) Prevention of premature capacitation, i.e. readiness to perform the acrosome reaction, where decapacitation factors, which are presumably epididymal polypeptides, play a regulating role.

10 (c) Change in the surface antigens of the spermatozoa in order to promote binding between the spermatozoa and oocyte.

(d) Change in the spermatozoal membrane in order to facilitate fusion with the ovum.

15 The metabolic event in the epididymis which induces these processes is largely unclear even in the animals studied. However, with the aid of gel electrophoresis it has been shown that some polypeptides are present in the epididymis, in addition to the seminal fluid originating from the rete testis. Among these, the polypeptides (a) to (e) of Cooper which contain so-called "acidic epididymal glycoprotein (AEG)" (Lea et al., (1978)) and polypeptides B to E of Brooks and colleagues (D.E. Brooks and J. Higgins, Characterization and androgen-dependence of proteins associated with luminal fluid and spermatozoa in the rat epididymis, Journal of Reproduction and Fertility 59, 363-375, (1980)) are present in the rat. The results of these 20 models indicate that the formation of these polypeptides is influenced by androgens.

25

30 However, the findings obtained from animal studies cannot be applied to humans because of the high tissue and species specificity. For example, the epididymal polypeptides essential in the rat or the mRNAs coding them have not been found in any other tissue and, apart from the mouse, in any other

species (D.E. Brooks et al., *Europ. J. Biochem.*, 161, 13-18 (1986); *J. Biolog. Chem.*, 261, 4956-4961, (1986)).

EP-A-0 440 321 discloses 5 specific polypeptides of the human epididymis and  
5 nucleotide sequences which code for these polypeptides which are closely  
connected with maturation of spermatozoa in the epididymis. These nucleotide  
sequences and the corresponding expression products, as well as antibodies  
directed against them, are suitable for diagnosis and use in the treatment of  
male infertility.

10

#### SUMMARY OF THE INVENTION

It is an object of the present invention to provide methods of diagnosis and  
treatment of male infertility, such as when caused by protein metabolism  
disturbances in the epididymis and compounds and compositions to perform  
these methods.

15

In one embodiment, the present invention provides epididymis-specific  
receptor protein (ESRP), represented as SEQ ID NO: 2, and active derivatives  
or protein fragments thereof having the same biological activity and/or  
immunogenicity, and DNA sequences which code for the ESRP in particular those  
according to SEQ ID NO: 1 and DNA sequences derived from these, DNA which  
code for the protein fragments, and, taking into consideration the  
degeneration of the genetic code, DNA sequences which coincide with these DNA  
sequences are within the scope of the present invention. (Osteroff, C;  
20 Ivell, R, Kirchhoff, C, *DNA Cell Biol* (1997) Apr; 16(4):379-389.)

25

In another embodiment, the present invention provides homologous amino acid  
sequences and corresponding DNA sequences having a degree of homology or  
similarity of at least about 70%, preferably at least about 80%, and  
30 particularly preferably at least about 90%. Finally, the invention also  
relates to nucleotide sequences which hybridize with sequences described

herein, under the conditions defined below and are suitable, in particular, as probes.

Prior to the present invention, it had not been possible to isolate and  
5 identify epididymis-specific receptor proteins participating in production of  
the epididymis medium described above, by conventional methods, because of  
the small quantities of tissue available.

It has now been possible, for the first time, to detect and characterize a  
10 specific transmembrane receptor protein of the mammalian epididymis and to  
enable its preparation by means of recombinant processes or chemical  
synthesis.

To identify the receptor protein according to the invention, a cDNA library  
15 of human epididymal mRNA in lambda gt11 (Clontech, California) was first  
established, and the epididymis-specific recombinants were then selected in a  
sequence of differential hybridization steps (cf. figures 1 and 3 to 5) and  
then subcloned in the bacterial plasmid pBS (Stratagene, California, USA) and  
characterized.

20 The aim of the differential screening was to isolate clones of mRNAs which  
are present in the tissue of the epididymis but not in the tissue of other  
organs and organs of related function. In the process employed according to  
the invention, human testicular tissue was used as the comparison material in  
25 the first step. Since some of the tissue samples used as the starting  
material originated from orchiectomized men who had been treated with various  
medicaments, it was first investigated whether and to what extent the  
medication has an influence on the particular polypeptide expression pattern.  
However, it was demonstrated on 7 patients of different age and different  
30 medication (cf. example 2 and figure 2a to c) that the patterns are  
influenced only insignificantly by prior administration of medicaments, and

also that the individual deviations are slight.

As shown in the diagram in figure 1, screening was carried out according to the invention in a primary and a secondary stage, potentially epididymis-specific clones being isolated in the primary stage by absence of cross-hybridization with testis mRNA, and these being selected further in the secondary stage by absence of cross-hybridization with mRNA from the brain and liver (cf. example 3 and figures 3 and 4).

10 By subsequent northern blot analysis against total mRNA from testes and human decidua, the number of possible clones could be concentrated further, and finally assigned to six independent cDNA families (cf. example 3 and figure 5). The cDNA clone, according to SEQ ID NO: 1, coding for the receptor protein according to the invention was isolated from one of these cDNA families.

15 The results obtained according to the invention show that the mRNA on which this clone is based is not synthesized in other human tissues (cf. figure 6); it is accordingly an epididymis-specific molecule. Figures 6 and 7 show the underlying autoradiograms of Northern blot hybridization in each case.

20 The nucleotide sequence and the corresponding amino acid sequence of the cDNA clone according to the invention were analysed as described in example 7. The result is shown in SEQ ID NO: 1.

25

The specificity of this clone and therefore of the corresponding expression product can be seen from its exclusive northern hybridization with RNA extracted from the human epididymis, which could not be detected by cross-hybridization in any other human tissue studied, including the organs of the 30 brain and liver. The tissue specificity was furthermore confirmed by the fact that a homologous gene probe produced from the rat using PCR technology

also hybridized exclusively with RNA from the epididymal tissue, but not with RNA from another tissue of this animal species (cf. figure 7).

Complete sequencing of a series of homologous cDNA clones isolated from two 5 human epididymal cDNA libraries and determined using the 5' RACE method (M.A. Frohman, RACE: Rapid Amplification of cDNA ends, in: PCR Protocols: A Guide to Methods and Applications, Academic Press Inc., p. 28 - 38 (1990)) gave a uniform gene transcript according to SEQ ID NO: 1 with a length of 4,665 nucleotides, which corresponds to the length of the mRNA determined by 10 northern hybridization. This sequence comprises a complete reading frame which codes for a polypeptide of 1,038 amino acids (SEQ ID NO: 2).

The homology or similarity of the amino acid sequences and DNA sequences according to the invention (SEQ ID NO: 1 and 2) with the sequences stored in 15 various international databanks (NIH GenBank database, EMBL, PIR) was analysed. It was found that the maximum homology found was with the "secretin/VIP superfamily" of G protein-coupled receptors (W.C. Probst et al., Sequence Alignment of G-Protein Coupled Receptor Superfamily, DNA and Cell Biology, 11, p. 1 - 20 (1992)) and was 25%.

20 The present invention, therefore, provides a hitherto unknown receptor, the long N-terminal end of which indicates an extracellular domain comparable to that of glycoprotein receptors, i.e. a new receptor for a large glycoprotein ligand originating from testicular secretion. With the aid of a 25 hydrophobicity analysis (example 7), seven clearly defined hydrophobic regions were demonstrated, which are to be designated transmembrane domains (cf. figure 8a). The structure of the protein according to the invention characterizes this as a membrane-resident receptor which is to be assigned to the gene family of G protein-coupled receptors (see above). It has the G 30 protein binding domains typical of this group of proteins, with the aid of which information is transduced through the receptor and passed on to the

inside of cells via G proteins (L. Birnbaumer, J. Abramowitz and A.M. Brown, Receptor-effector coupling by G proteins, *Biochem. Biophys. Act.*, 1031, 163 - 224 (1990)).

5 The receptor protein according to the invention is present in the epididymis in a high abundance (the mRNA makes up about 0.01% of the cDNA library) and *in situ* hybridization studies localized the mRNA in the epithelial cells which line the ductus epididymis (figure 9a).

10 Furthermore, the receptor protein according to the present invention, is a molecule which is highly conserved among mammalian species, since in the mammalian species studied to date, homology of the expression products of about 90% was found, with the same tissue specificity (figure 7 and 9c, d).

15 The nucleotide sequences which code for the new receptor protein and fragments according to the invention can be transferred into prokaryotic or eukaryotic host cells by conventional methods via suitable vectors and expressed there as protein. The present invention includes vectors and host cells containing these nucleotide sequences as well as host cells transformed by same, and the recombinantly expressed proteins and fragments.

20 All the DNA sequences which, after transformation of suitable prokaryotic and/or eukaryotic host cells, ensure the production of nucleic acids for use as diagnostics and/or the expression of proteins or polypeptides which have at least a portion of the primary structure and one or more of the biological and/or immunogenic properties of the receptor protein according to the invention are suitable and included according to the invention. These sequences, in single- or double-stranded form, include, in particular:

30 a) The nucleotide sequence represented as SEQ ID NO: 1, the sequence of nucleotides 1 to 3,114 of SEQ ID NO: 1, nucleotide, sequences which

are homologous to the abovementioned sequences with a degree of homology or similarity of at least about 70%, and fragments thereof, where these fragments code for polypeptides or proteins having the same biological activity and/or immunogenicity, and syngenetic or complementary sequences;

b) nucleotide sequences which hybridize with the protein-coding region of the nucleotide sequence described under a), for example under the hybridization conditions described in examples 3 or 4;

c) nucleotide sequences which, with the exception of the deviations caused by the degeneration of the genetic code, hybridize with the sequences mentioned under a) and/or b); and

d) nucleotide sequences which code for the amino acid sequence represented as SEQ ID NO: 2, sequences which are at least about 70% similar, at least about 80% similar, or at least about 90% similar to nucleotide sequences coding for the amino acid sequence represented as SEQ ID NO: 2, and fragments thereof which code for polypeptides or proteins which have at least one biological activity and/or immunogenicity as the protein represented by SEQ ID NO: 2, and syngenetic or complementary sequences.

The term "syngenetic sequence" includes all sequences which are derived from the same or a homologous or a similar gene and code for the receptor protein in the context of the invention, or can be used for the preparation of probes. The term also includes, in particular, sequences which show deviations on the basis of degeneration of the genetic code, as well as RNA sequences.

Examples of suitable modifications of the DNA sequence are nucleotide

substitutions which do not give rise to another amino acid sequence of the ESRP, but which may correspond to the codon usage of the host organism into which the DNA construct is introduced or nucleotide substitutions which do give rise to a different amino acid sequence and therefore, 5 possibly, a different protein structure without, however, impairing the properties of the native variant. Other examples of possible modifications are insertion of one or several nucleotides into the sequence, addition of one or several nucleotides at either end of the sequence, or deletion of one or several nucleotides at either end or within the sequence.

10

15 The invention also relates to derivatives in the form of naturally occurring allelic variations of the receptor protein according to the invention or fragments of such allele variations, it being possible for the various allelic forms to differ in respect of sequence length and in respect of deletions, substitutions, insertions or additions of amino acids, both in each case from one another and from the amino acid sequence described in SEQ ID NO: 2.

20 According to the invention, the term "proteins or polypeptides having the same biological activity and/or immunogenicity" designates molecules which have a) the same epididymal specificity and b) the same ligand-binding capacity as the proteins and polypeptides identified according to the invention. According to the invention, the term "ligand" includes both 25 antibodies against any desired epitopes included in the proteins or polypeptides according to the invention, and other chemical substances or molecules which are capable of binding to one or more of the domains present in the proteins or polypeptides according to the invention. In one embodiment, the present invention provides isolated or purified peptides N1, 30 N2, A, B and C as described in Table I.

Suitable vectors for prokaryotic host cells according to the present invention are, for example, plasmids of the pET series (Rosenburg et al., Vectors for selective expression of cloned DNAs by T7 RNA Polymerase, Gene, 56, p. 125-135 (1987)), of the pGEX series (Pharmacia, Freiburg), of the pRIT series (Pharmacia, Freiburg) and of the pH6EX3 series (Berthold et al., Purification of Recombinant Antigenic Epitopes of the Human 68-kDa (U1) Ribonucleoprotein Antigen Using the Expression System pH6EX3 Followed by Metal Chelating Affinity Chromatography, Protein Expression and Purification, 3, p. 50-56 (1992)). Examples of prokaryotic host cells according to the 10 present invention include laboratory strains of Escherichia coli K12 and the strains BL21 (DE3) and LE392.

Examples of suitable vector systems according to the present invention for eukaryotic host cells in the case of mammalian cells, are SV40 viruses, polyoma viruses, adenoviruses, various retroviruses, papilloma viruses (P.W.J. Rigby, Expression of cloned genes in eukaryotic cells using vector systems derived from viral replicons, Genetic Engineering. vol. 3, p. 84-141 (1982)) and vaccinia viruses (Mackett et al., The construction and characterisation of Vaccinia Virus recombinants expressing foreign genes, DNA 15 Cloning, vol. II. (1985), IRL Press, Oxford), and derivatives thereof, as well as plasmids which contain parts of viral genes (e.g. pSV2) and can be employed as "shuttle vectors" (P.W.J. Rigby). Possible vectors for insect cells are, for example, pJVETL baculoviruses (Invitrogen, San Diego, Calif., USA), while for yeast cells, e.g., pJP31 and YEp- and YIp-plasmids (Carter et al., Expression and secretion of foreign genes in yeast. DNA Cloning, vol. 20 III, (1987), IRL Press, Oxford) can be used.

Suitable mammalian cells are e.g. COS cells, CHO cells, AtT20 cells, NIH3T3 cells (Rigby, Mackett et al., HEK293 cells (American Type Cell Collection 30 (ATCC) No. CRL 1573) and MDCK cells (European Collection of Animal Cell Cultures No. 85011435), while suitable insect cells are e.g. Sf9, Sf21 and

TN5 (Invitrogen, San Diego, Calif., USA). Suitable yeast cells are e.g. X4003-5B (Carter).

The receptor protein according to the invention can furthermore be  
5 synthesized chemically by known processes (J.M. Stewart, Synthesis and use of  
neuropeptides, Neuroendocrine Peptide Methodology, ed. P.M. Conn, Academic  
Press, New York, p. 815-844 (1989)). The same applies to polypeptides or  
peptide epitopes having the same immunogenicity which are coded by fragments  
or syngenic sequences of the DNA sequence according to the invention (SEQ ID  
10 NO: 1). This is also true for the oligopeptides N1 (SEQ ID NO: 3), N2 (SEQ  
ID NO: 4), A (SEQ ID NO: 5), B (SEQ ID NO: 6) and C (SEQ ID NO: 7) (table I).

The polypeptides and proteins according to the invention, the nucleotide  
sequences which code for these, including their complementary sequences, and  
15 antibodies produced on the basis of the polypeptides and proteins offer for  
the first time the possibility of diagnosing and, where appropriate, treating  
disturbances in the protein metabolism of the epididymal epithelium, and of  
providing new contraceptive agents.

20 Thus, for example, the above mentioned nucleotide sequences of the present  
invention can be provided with markers and used as probes for *in situ*  
hybridization in tissue diagnostics of biopsy samples or thin sections, in  
order to determine the physiological state of the tissue in respect of the  
presence and concentration of the receptor proteins according to the  
25 invention, and to compare it with standard values.

30 Polyclonal and monoclonal antibodies for use in immunological detection  
methods can be produced in a known manner with the aid of the highly pure  
polypeptide according to the invention. Such antibodies can be produced on  
the basis of the complete receptor protein and on the basis of fragments and  
active derivatives thereof, where these have the same immunogenicity (cf.

examples 9).

The antibodies can be marked or labeled and used *in vitro* or *in vivo* for detection of the receptor protein according to the invention.

5

The receptor protein according to the invention and biologically active derivatives or fragments thereof having the same immunogenicity can furthermore be used in marked or labeled or non-marked or unlabeled form as antigens for identification of autoantibodies in the sera of infertile men.

10 This possibility is of particular importance, since it is assumed that in a large proportion of cases infertility is to be attributed to the presence of autoantibodies against essential components of the reproductive system. However, the test methods available to date measure only antibodies directed against some sperm surface antigens, and a sufficiently high titre of the antibodies must be present to allow sperm agglutination to take place. It is assumed, however, that antibodies are present in far lower titres and can cause infertility. These can be detected as an antigen with the purified or isolated receptor protein prepared according to the invention.

20 Starting from the amino acid sequence disclosed according to the invention, two different processes are available for isolation of antigens for the production of antibodies.

Firstly, a potentially immunogenic region of the protein sequence of 25 interest, which on the one hand is relatively hydrophilic and therefore lies on the outside of the protein molecule, but on the other hand is not impaired in its steric conformation by formation of cysteine disulfide bridges or possible glycosylation sites, can be selected with the aid of a computer.

30 This peptide region is then synthesized, if appropriate together with flanking amino acids, subsequently coupled to carrier substances and employed

as an immunogen for the induction of antibodies (cf. example 9).

Alternatively, for example, the nucleotide sequence which codes for the polypeptide of interest can be cloned into a suitable expression plasmid 5 vector. After subsequent transformation into suitable bacteria, these vectors allow an inducible expression of the coded polypeptide. The bacteriogenic protein or protein fragment prepared in this manner can be used directly, after purification from the bacterial extract, as an antigen for immunization for induction of antibodies (cf. example 8).

10

The antibodies can be provided with a detectable marker, such as e.g. a fluorescent molecule (fluorophor) and can be used, for example, in tissue samples to determine the presence of the epididymis-specific receptor protein in the epididymal epithelium with the aid of immunofluorescence.

15

The identification and characterization of the receptor protein according to the invention as a highly specific mediator molecule which is present exclusively on the cells of the epididymal epithelium of mammals and is capable of transmitting information for control of the cell function within 20 the cells of the epithelium make it an extremely interesting candidate for diagnosis and, if appropriate, influencing of the abovementioned physiology of the epididymal epithelium. The present invention provides therefore, a therapeutic method and composition to improve the spermatozoa maturation process in the epididymis, for example by administration of the ligand 25 lacking or formed in an inadequate amount in the individual to be treated, in a pharmaceutically acceptable carrier or diluent.

The present invention also provides a contraceptive method and compositions 30 which have a negative influence on the epithelium which can lead to deterioration of spermatozoa maturation. A synthetic ligand which indeed binds firmly to the receptor protein according to the invention but induces

no signal transfer or transmission is suitable, for example, for this purpose. Antibodies directed against the receptor protein according to the invention or active derivatives or fragments thereof having the same immunogenicity can furthermore be employed in order to impede binding of and therefore signal inducement by the ligand or ligands specific to the receptor protein by way of competitive displacement and are also a part of the present invention.

Using certain methods, such as e.g. phage display and peptide display (J.K.

10 Scott and G.P. Smith, Searching for peptide ligand with an epitope library, Science 249, p. 386-390 (1990)); J.J. Devlin et al., Random peptide libraries: a source of specific protein binding molecules, Science 249, p. 404-406 (1990)), Evolutive Biotechnology (M. Eigen, Selforganization of matter and the evolution of biological macromolecules, Die Naturwissenschaften 58, p. 465-523 (1971); M. Eigen, Automated molecular evolution, Max-Planck Institute of Biophysical Chemistry (1991)), one is able to identify synthetic ligands which have the ability to bind specifically and with a high affinity to such a receptor protein and to act either agonistically or antagonistically on the signal transfer capability thereof (ex. 11). Molecules by means of which the physiology of cells which express the receptor protein can be influenced positively (therapeutically) or negatively (contraceptively) can be provided by this route.

Alternatively, the ESRP of the invention, expressed in a host cell, as described herein, may be expressed to retain the transmembrane and, optionally, the cytoplasmic region of the native variant, to be anchored in the membrane of the host cell, and the cells carrying the ESRP may be used as such in the screening or diagnostic assay. Alternatively, the receptor may be a component of membrane preparations, e.g. in solubilised and/or reconstituted form.

The ESRP, derivative or analogue of the invention may be immobilized on a solid support and may, as such, be used as a reagent in the screening methods of the invention. The ESRP, derivative or analogue may be used in 5 membrane-bound form, i.e. bound to whole cells or as a component of membrane preparations immobilised on a solid support.

The solid support employed in the screening methods of the invention preferably comprises a polymer. The support may in itself be composed of 10 the polymer or may be composed of a matrix coated with the polymer. The matrix may be of any suitable material such as glass, paper or plastic. The polymer may be selected from the group consisting of a plastic (e.g. latex, a polystyrene, polyvinylchloride, polyurethane, polyacrylamide, polyvinylalcohol, nylon, polyvinylacetate, and any suitable copolymer 15 thereof), cellulose (e.g. various types of paper, such as nitrocellulose paper and the like), a silicon polymer (e.g. siloxane), a polysaccharide (e.g. agarose or dextran), an ion exchange resin (e.g. conventional anion or cation exchange resins), a polypeptide such as polylysine, or a ceramic material such as glass (e.g. controlled pore glass). 20

The physical shape of the solid support is not critical, although some shapes may be more convenient than others for the present purpose. Thus, the solid support may be in the shape of a plate, e.g. a thin layer or microtiter plate, or a film, strip, membrane (e.g. a nylon membrane or a 25 cellulose filter) or solid particles (e.g. latex beads or dextran or agarose beads).

It is furthermore contemplated to locate the ligand-binding site on the ESRP of the invention, for instance by preparing deletion or substitution 30 derivatives of the native ESRP (as described herein) and incubating these with ligands known to bind the full-length ESRP and detecting any binding

of the ligand to the ESRP deletion derivative. Once the ligand-binding site has been located, this may be used to acquire further information about the three-dimensional structure of the ligand-binding site. Such three-dimensional structures may, for instance, be established by means of 5 protein engineering, computer modelling, NMR technology and/or crystallographic techniques. Based on the three-dimensional structure of the ligand-binding site, it may be possible to design substances which are agonists or antagonist to the ESRP molecule.

10 The present invention is further explained in detail below with the aid of examples, figures and sequence protocols.

Brief description of the figures

**Figure 1**

Schematic representation of the a) primary and b) secondary screening for the 20 receptor protein according to the invention.

**Figure 2**

25 Two-dimensional electrophoresis patterns of translation products derived from epididymis- or testis-specific poly(A)<sup>+</sup> RNA and synthesized in a cell-free system in the presence of [<sup>35</sup>S]-methionine.

30 a. epididymal tissue of a 74 year-old orchiectomy patient treated with cyproterone acetate;

b. epididymal tissue of a 58 year-old orchiectomy patient not treated with medicaments;

c. testis tissue of the same patient not treated with medicaments (see figure 5 2b).

**Figure 3**

10 Autoradiograms of two replica filters from one of the Petri dishes which were hybridized during the primary screening with

a. complex epididymal cDNA probes or

15 b. complex testicular cDNA probes.

**Figure 4**

20 Autoradiograms of three replica filters from one of the Petri dishes which were hybridized with complex epididymal cDNA probes, complex liver cDNA probes or complex brain cDNA probes during the secondary screening. ESPR-positive clone is indicated coding for the receptor protein of the invention.

25 **Figure 5**

Examples of autoradiograms of the northern blot analyses of potentially positive cDNA clones, E = epididymis, T = testis, D = decidua

30 a. Positive cDNA clone, indicated by epididymis specific hybridization;

- b. Positive but not epididymis-specific hybridization pattern;
- c. Positive cDNA clone (\*), which also hybridizes with rRNA;
- 5 d. cDNA clone without tissue-specific expression.

**Figure 6**

10 Northern blot analysis of human tissues. Blots carrying 20 :g of total RNA per lane [except poly(A)<sup>+</sup>epididymis, and vas deferens lanes which only contained 10 :g of RNA] of various tissues were hybridized with different fragments of the receptor protein of the invention (compare figure 8b, bold lines).

15 a. 1.4-kb 3'UTR

b. 0.7-kb 5'Bam HI

20 c. 1-kb 5'RACE.

Probes were prepared by random primer extension using  $["^{32}\text{P}]\text{dCTP}$  (Sambrook et al., "Molecular Cloning - A Laboratory Manual", 2nd edition, Cold Spring Harbor Laboratory Press (1989)).

25

**Figure 7**

Northern blot analysis of rat tissues. Blots carrying 20 :g of total RNA per lane (except for lane 2 containing only 4 :g of RNA) of different rat tissues were hybridized with the E-F1 fragment of the rat homologue of the receptor

protein (compare figure 8b, PCR subclones). The homologous probe was prepared by random primer extension using ["<sup>32</sup>P]dCTP (upper and lower panels). The upper blot was subsequently hybridized to an actin probe to show equal loading (middle panel).

5

**Figure 8**

Diagrammatic representation of receptor protein cDNA

10

a. Hydrophilicity plot of the coding region containing a hydrophobic signal peptide (amino acids 25-62) and seven hydrophobic amino acid stretches between amino acids 642-915 (numbers according to SEQ ID NO: 2).

15 b. Putative structure of the whole known sequence of the receptor protein of the invention and description of isolated cDNA clones.

The predicted coding region (1,038 amino acids) is given as a box, the shaded bars standing for the seven transmembrane domains (I-VII) corresponding to the hydrophilicity plot given in a.

20 3' UTR, originally isolated clone from the first 8gt11 epididymal library (library 1);

fragment "virtually complete in length", clones isolated from the 8 Uni-ZAP library (library 2);

25 5' Bam HI and 3' Kpn I, restriction fragment subclones of the 5' and 3' regions;

PCR subclones are the 5' RACE fragments derived from nested primers (A1, A2) from the 5' region of the 3.7-kb cDNA fragment "virtually complete in length" using total human epididymal cDNA, and those derived from primer pair E and F1 using epididymal cDNA from human, rat, mouse, and dog. Bold lines indicate 30 fragments used as probes in other figures.

(A)<sub>n</sub>, poly(A) tail; C, carboxyl terminus; kb, kilobases; N, amino terminus;

nt, nucleotides; SP, putative signal peptide.

c. Structure of receptors with seven helices

(a) Linear diagram of the receptor protein. Seven blocks of hydrophobic amino acids (1 - 7) are shown. Each assume the shape of a helix, which presumably extends through the membrane. The helices are connected by loops which lie alternately outside (E) and inside (C) the cell. The N-terminus of the receptors points towards the outside, while the C-terminus points inside the cell.

10 (b) Model of the assumed structure of the folded receptor.

**Figure 9**

15 *In situ* localization of receptor protein transcripts in the human and rat epididymis.

a. human caput region

20 b. human caput region

c. rat epididymis (proximal)

25 d. rat epididymis (distal)

Tissue cryosections were hybridized either with  $^{35}$ S-labeled antisense cRNA fragment E-F1 of the receptor protein (a), or with  $^{35}$ S-labeled sense control cRNA (b). The rat tissue was hybridized with the equivalent rat-specific 30 antisense cRNA probe (c,d). The hybridizing transcripts were visualized as white grains on dark-field reflectance micrographs.

**Figure 10**

5 Immunohistochemical staining of frozen tissue sections derived from

a. human epididymis

b. human epididymis

10

c. mouse epididymis

d. mouse epididymis

15

e. bovine epididymis, and

f. bovine epididymis.

In Figures 10 a, c and e anti-B antiserum was employed.

20

In Figures 10 b, d and f pre-immune serum was employed (negative control).

**Examples**

**Example 1**

25 **RNA preparation**

Total RNA was extracted by the method of J.M. Chirgwin et al., Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease, Biochemistry, 18, 5294, (1979), using guanidine isothiocyanate, from frozen tissues of the human testis and epididymis which had been obtained by 30 orchiectomy from men aged 58 and 74 years suffering from prostate carcinoma, and was then purified by caesium chloride density gradient centrifugation.

Using an oligo(dT)-cellulose column, poly(A)<sup>+</sup>RNA could be enriched by affinity chromatography, as described by H. Aviv & P. Leder, Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose, Proc. Natl. Acad. Sci., USA, 69, 1408-1412, (1972). The RNA samples were precipitated with ethanol and, after resuspension in sterile water, stored in a concentration of 1 :g/:l at -80EC.

Example 2

In vitro translation and characterization of the products by two-dimensional gel electrophoresis

0.5 to 1 :g of the poly(A)<sup>+</sup>RNA of the epididymis and testis from example 1 was in each case translated in vitro in a cell-free system of a reticulocyte lysate from the rabbit (New England Nuclear (NEN), Dreieich, FRG). The synthesis was carried out in the presence of (<sup>35</sup>S)-methionine (specific activity > 1,000 Ci/mmol). The resulting protein products were then separated electrophoretically in a two-dimensional gel (cf. P.H. O'Farrel, High resolution two-dimensional electrophoresis of proteins, J. Biol. Chem., 250, 4007-4021 (1975)). For the first dimension, an amount of 350,000 cpm of the (<sup>35</sup>S)-methionine-labelled translation batches was in each case applied to the gel and focused isoelectrically at an applied voltage of 10,000 V.h. The pH gradient was between 4.0 and 7.5. The separation of the polypeptides according to their molecular weight was carried out in the second dimension within a linear gradient of 7.5 - 15% acrylamide (cf. D.M. Neville & H. Glossmann, Molecular weight determination of membrane protein and glycoprotein subunits by discontinuous gel electrophoresis in dodecylsulphate, Meth. Enzym., 32, 92-102, (1974)). To estimate the molecular weights, a <sup>14</sup>C-labelled protein marker (Amersham, GB) was applied in the second dimension. The gel was then subjected to fluorography with the aid of Kodak X-AR5 autoradiographic film (cf. W.M. Bonner & R.A. Laskey, A film detection method for tritium-labelled proteins and nucleic acids in

polyacrylamide gels, *Europ. J. Biochem.*, 46, 83-88, (1974)). The exposure time was 8 days. The results are shown in figure 2.

Figure 2, which shows the results, shows the pattern of the translation products derived from epididymis- or testis-specific poly(A)<sup>+</sup>RNA and synthesized in a cell-free system.

In particular, the figure shows

10 2a: epididymal tissue of a 74 year-old orchiectomy patient treated with cyproterone acetate,

15 2b: epididymal tissue of a 58 year-old orchiectomy patient not treated with medicaments,

20 2c: testis tissue of the same patient not treated with medicaments (see figure 2b).

Epididymis-specific translation products of which the corresponding bands do not appear in figure 2c (translation products of mRNA from the testis) are identified by small arrowheads (see figure 2b).

25 A comparison of figures 2a and 2b shows a change in the expression in epididymal tissue due to previous treatment with antiandrogens. Bands which are reduced in figure 2a compared with figure 2b have been marked by large arrowheads. The molecular weights of the protein marker are given in kilodalton. Actin(A)- and tubulin(T)-like products are identified in each case.

30 Example 3

Compiling the cDNA complete library and selection of potentially specific clones

The bacteria and bacteriophages were cultured, handled and obtained and the DNA recombinant techniques were used as described by T. Maniatis et al.,  
5 Molecular Cloning, A Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbour, N.Y., (1982).

10 a) Poly(A)<sup>+</sup>RNA from the epididymal tissue of a patient who had not been treated with medicaments was used for compiling the cDNA library (with regard to the pattern of the in vitro translation products, cf. figure 2b). 20 :g poly(A)<sup>+</sup>RNA according to example 1 were subjected to reverse transcription using oligo(dT) as a primer (cf. U. Gubler & B.J. Hoffmann, A simple and very efficient method for generating cDNA libraries, Gene, 15, 263-269, (1983)).

15 The double-stranded cDNA constructs were ligated on both sides with EcoRI-linkers and then inserted into the EcoRI cloning site of the bacteriophage lambda gt11 (Clontech California).

20 To discover clones of complete length, another epididymis-specific cDNA library was compiled substantially by the method described, but with the deviation that the cDNA constructs were cloned unidirectionally. This was effected by ligating the particular 5' end with an EcoRI-linker and the particular 3' end with a Xhol-linker and then inserting the particular constructs into the bacteriophage lambda  
25 Uni-ZAP (Stratagene, La Jolla, California).

30 Bacteria of E. coli strains Y 1090 and XL1-Blue were cultured to the logarithmic phase, transferred to soft agarose, plated out together with the non-amplified lambda library in a phage density of 500 to 1,000 pfu per Petri dish (15 cm) and incubated for 8 hours at 42EC or

37EC.

For the differential screening, the plaques of each dish were subsequently replica-plated on two nitrocellulose filters (Schleicher & Schuell, Darmstadt, FRG, BA85), the incubation time for the first filter being 1 minute and that for the second filter being 2 minutes.

Replica filters of a total of 20 Petri dishes were prepared in this manner and then hybridized (positive/negative) with in each case two single-stranded radioactively labelled cDNA probes (cf. b, below).

b) The probes of poly(A)<sup>+</sup>RNA from the epididymis (gives a positive probe) and poly(A)<sup>+</sup>RNA from the testis (gives a negative probe) were prepared as follows using oligo(dT) as the primer:

The two RNA species originated from the corresponding tissues of a patient who had not been exposed to prior treatment with medicaments.

In each case 1 :g poly(A)<sup>+</sup>RNA according to example 1 was denatured for 5 minutes at 65EC in a volume of 2 :1. After rapid cooling of the batches on ice, in each case the following constituents were added in succession:

- 2 :1 oligo(dT) (100 :g/ml)
- 1 :1 dNTP-Mix (dATP, dGTP, dTTP, in each case 2 mM)
- 1 :1 40 mM sodium pyrophosphate
- 1 :1 RNasin (Amersham, GB)
- 2 :1 10x reverse transcriptase buffer (500 mM Tris-Cl (pH 8.5), 500 mM KCl, 100 mM MgCl<sub>2</sub>, 100 :g/ml BSA, 10 mM EDTA, 10 mM dithiothreitol)
- 20 units AMV reverse transcriptase (Boehringer Mannheim, FRG)
- 10 :1 ("<sup>32</sup>P)-dCTP (NEN, 10 :Ci/:l; S.A. > 3,000 Ci/mmol).

The reaction batches were incubated for 15 minutes at 42EC. After addition of 1 :1 "Chasemix" (containing all four dNTPs in a concentration of 10 mM each), the incubation was continued for a further 20 minutes. The reaction 5 was stopped by addition of in each case 1 :1 0.5 M EDTA and the products were precipitated with ethanol without further purification. The probes, marked with a specific activity of  $> 10^8$  cpm/:g, were hybridized in parallel with the abovementioned replica filters. The hybridization was carried out in 5x Denhardt's solution, 4x SET (200 mM Tris (pH 8.0), 20 mM EDTA, 0.6 M NaCl), 10 0.1% sodium pyrophosphate and 25 mM sodium phosphate buffer (pH 7.0) for 72 hours at 65EC and a concentration of the radioactivity of  $5 \times 10^6$  cpm/ml. The filters were then washed in 0.1% SDS, 2x SSC (300 mM sodium chloride, 30 mM sodium<sub>3</sub> citrate) at a temperature of 65EC.

15 Figure 3 shows the autoradiograms of two replica filters from one of the Petri dishes which were hybridized with epididymal (a) or testicular (b) cDNA probes. The autoradiograms were developed after an exposure time of 16 hours using an amplifying film. The filters represent about 500 independent clones 20 of the cDNA total library from epididymal tissue compiled in the lambda gt11 system. The positive epididymis-specific hybridization signals are marked with arrowheads (cf. figure 3a).

With the aid of this primary screening method (cf. figure 1a), it was possible to analyse about 10,000 independent cDNA clones. 265 recombinants 25 of which the signals after hybridization with the epididymal cDNA probe were very much more intense compared with those prepared from testicular tissue were identified. Based on the portion of the phage library analysed by screening, the number of positive cDNA clones corresponds to a proportion of 2.5%.

30

Positive cDNA clones were isolated and divided into groups of up to 80 clones

for secondary screening.

The process described above was repeated, with the deviation that in each case 3 replica filters were prepared from each of the Petri dishes covered by 5 groups of in each case 80 clones. To prepare the negative cDNA probes, poly(A)<sup>+</sup>RNA from the human brain or from the human liver (Clontech California, USA) was used here.

From the analysis of this secondary screening process (cf. figure 1b), the 10 number of positive recombinants was reduced to 59 clones (cf. figure 4).

Positive clones were isolated without further purification of the plaques. The purified lambda DNA was cut with EcoRI and the inserts were isolated by Biotrap elution (Schleicher & Schuell). The EcoRI inserts were subcloned into 15 the bacterial plasmid vector pBS (Stratagene, California, USA) and inserted by the CaCl<sub>2</sub> transformation method into the bacteria cells of the E. coli strain XL1-Blue (Stratagene, California) (cf. T. Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbour, N.Y., (1982)).

Example 4

Preparation of northern blots for analysis of the tissue-specific gene expression and classification of the epididymis-specific cDNA clones in families with related sequences

20 :g total RNA from the human epididymis (E) of a patient treated with cyproterone acetate and the same amount of total RNA from the human testis (T) and human decidua (D) were separated electrophoretically in a horizontal 1.3% formaldehyde-agarose gel for 16 hours at a constant voltage of 25 volt 30 (cf. Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbour, N.Y., (1982) (figures 5, 6 and 7)).

To determine the size of the RNAs, the gel was additionally loaded with an RNA marker (0.24-9.5 kB). The RNA was then transferred to nylon membranes (Hybond N, Amersham) by capillary blotting in the presence of 20 x SSC. The blots were subsequently hybridized successively with radioactive potentially 5 epididymis-specific cDNA probes. The probes were isolated by EcoRI restriction of the lambda gt11 DNA subcloned in the plasmid pBS and subsequent Biotrap elution, and after radioactive labelling had a specific radioactivity of  $> 10^9$  cpm/:g (cf. A.P. Feinberg and B. Vogelstein, A technique for radiolabeling DNA restriction endonuclease fragments to high 10 specific activity, *Analytical Biochemistry*, 132, 6-13, (1983)). The hybridization was carried out overnight at a radioactivity concentration of 1-2  $\times 10^6$  cpm/ml. The filters were subsequently washed stepwise with an increasing stringency, starting with 2 x SSC at room temperature up to 0.1 x SSC at a temperature of 65EC. For dehybridization, the filters were 15 incubated for 15 minutes at 65EC in the presence of 2 mM Tris (pH 7.5), 1 mM EDTA and 0.1% SDS. The filters were analysed autoradiographically for the presence of probe material, before they were used for a repeat hybridization with another probe if signals were absent.

20 Figure 5 shows examples of the autoradiograms of the northern blot analyses of potentially positive cDNA clones. The exposure time was 20 hours. The clones were evaluated qualitatively in respect of their tissue specificity from the various hybridization patterns.

25 a) Positive cDNA clone, indicated by epididymis-specific hybridization.

b) Positive but not epididymis-specific hybridization pattern.

30 c) Positive cDNA clone(\*), which also hybridizes with rRNA.

d) cDNA clone without tissue-specific expression.

A number of 36 epididymal cDNA clones could be identified with the aid of this method. The specific hybridization signal of each clone insert with 5 epididymal RNA was at least one order of magnitude more intense than that with RNA from the testis or decidua (cf. figure 5a).

Using the results of cross-hybridizations with the original phage library, the 36 epididymis-specific cDNA clones could be assigned to six independent 10 families, each of which comprised a set of clones related to one another. It was found that five of the cDNA clone families are derived from different, relatively short mRNA molecules having an average length of 600 to 1,000 15 nucleotides, while the clone according to the invention, which contains the genetic information for the present receptor protein, has a length of about 5,000 nucleotides.

Since only cDNA fragments having a maximum length of 3.7 kb could also be isolated from the second cDNA library (cf. fig. 8b), but on the basis of the 20 results of the northern hybridization an mRNA length of about 5 kb was to be assumed, the 5' start of this cDNA was cloned using the 5' RACE method (RACE = rapid amplification of cDNA ends; M.A. Frohman et al., "Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer", Proc. Natl. Acad. Sci. USA, 85, p. 8998-9002 (1988)). For this, two gene-specific antisense primers A1 and A2 were 25 prepared from the 5' region of the known cDNA sequence comprising 3.7 kb (A1 = 5'AGC TAT GGG AGC TGA AG 3' (SEQ ID NO:8) A2 = 5'TGT CAA TGG CAG GGC TG 3' (SEQ ID NO: 9)). With the aid of these primers, starting from 1 :g total RNA from the human epididymis, a 5' fragment having a length of 996 nucleotides, which overlapped with the already known cDNA sequence over a length of 49 30 nucleotides (incl. A1 primer) was obtained using the protocol "5' RACE System for Rapid Amplification of cDNA Ends" from Gibco BRL, Berlin (cf. fig. 8b, 5'

RACE-fragment, 1.0 kb).

Example 5

Checking of the tissue specificity of the clone according to the invention

5

In each case 20 :g total RNA from the human decidua, epididymis and testis were separated and transferred to nylon membranes according to example 4. To check the tissue specificity of the clone according to the invention, in each case 10 :g poly(A)<sup>+</sup>RNA from the epididymis and total RNA from the seminal duct (vas deferens), and in each case 20 :g total RNA from a lymphoblastoid cell line (Daudi), from prostate carcinoma cell lines (LNCaP, DU145) and from embryonal, epididymal cell lines (RVP, REP) were also separated analogously and likewise applied to a nylon membrane. A <sup>32</sup>P-labelled cDNA insert from the 3' region of the clone according to the invention (cf. fig. 8b, 3' UTR) was employed as the probe in figure 6a, and was hybridized with northern blots according to example 4. In figure 6b, in contrast, a <sup>32</sup>P-labelled BamHI cDNA insert from the 5' region of the clone according to the invention was employed (cf. fig. 8b, 5' BamHI fragment). Figure 6c shows a northern blot after hybridization with a <sup>32</sup>P-labelled 5' RACE fragment of the clone according to the invention.

The results shown in figure 6a-c show that the cDNA clone according to the invention gives an unambiguous hybridization signal only with the epididymal RNA. Furthermore, with the aid of the autoradiogram corresponding to figure 25 6c, the genuineness of the 5' RACE fragment could be confirmed (note the exclusive signal with epididymal RNA at approx. 5 kb). No cross-hybridization was to be detected in other types of human tissue analyzed (16 in total).

Example 6

30 Detection of homologous sequences in the epididymis of other species of mammals and checking of the tissue specificity of the clone according to the

**invention in the rat**

To investigate whether homologues of the receptor protein according to the invention are also expressed with the same tissue specificity in other 5 species of mammals an attempt was made to amplify corresponding fragments from epididymal total RNA from the rat, mouse and dog with the human subfragment E - F1 (cf. fig. 8b, PCR subclones, prepared using the two primers E = 5'CAT CCG AAA ATA CAT CC 3' (SEQ ID NO:10) and F1 = 5'TGA AGG CAC ACA TCT CC 3' (SEQ ID NO:11)) using the PCR technique (gen. ref., cf. R.S. 10 Cha and W.G. Thilly, Specificity, Efficiency and Fidelity of PCR, PCR Methods and Applications, vol. 3, p. 18 - 29 (1993)).

In each case 5 :g total RNA from the animal species mentioned were subjected 15 to annealing in a total volume of in each case 20 :l using 0.5 :g oligo(dT)<sub>12-18</sub> (Pharmacia, Freiburg). The subsequent first strand synthesis was carried out in the presence of in each case 10 mM dATP, dGTP, dCTP and dTTP and of 200 units of reverse transcriptase (Superscript, Gibco BRL, Berlin). For the subsequent PCR amplification, in each case 1 :l of the products of the first strand synthesis were mixed with in each case 10 mM of the four dNTPs, in each case 20 pM of the human gene-specific primers E and F1 (see above) and with in each case 7 units of Taq-Polymerase (Promega, Heidelberg) in a total 20 volume of in each case 50 :l and subjected to the following PCR cycles: denaturing (5 minutes at 95EC); then 30 cycles comprising denaturing (1 minute at 95EC), annealing (1 minute at 55EC) and elongation (1 minute at 72EC). The complementary cDNA fragments obtained in this manner were 25 separated electrophoretically in 1.2% agarose gels. Fragments having a length of about 750 nucleotides were then electroeluted and inserted into the cloning vector pCRII (Invitrogen, ITC Biotechnology, Heidelberg). The sequence of the cloned DNA fragments was determined as described in the 30 following example 7 and showed - both at the nucleotide and at the amino acid level - a particular homology of about 90% to the human sequence. This

surprising result shows that the DNA sequence which codes for the receptor protein according to the invention has a high degree of conservation at least among mammals.

5 For control of the tissue specificity of the clone according to the invention in the rat, in each case 20 :g of total RNA from various rat tissues (with the exception of the 2nd lane in figure 7 (testis), only 4 :g) were separated and transferred to nylon membranes according to example 4. A <sup>32</sup>P-labelled cDNA insert (spec. activity 19<sup>9</sup> cpm/:g) from the rat homologue of the clone 10 according to the invention (E - F1) was employed as the probe and hybridized with northern blots according to example 4. The autoradiograms of the northern blot analysis for the clone according to the invention developed after an exposure time of 15 hours are shown in diagram form in figure 7. A rehybridization of the blot shown above with a control actin probe is shown 15 in the middle of figure 7, which shows that approximately equimolar amounts of RNA were found in all the lanes of the gel.

20 Figure 7 shows that probe E - F1 from the rat hybridized exclusively with epididymal RNA, the signals being the most intense within the epididymis in the caput (proximal region) and in the cauda (distal region) (cf. lanes 5-7).

The results in the rat confirm the epididymis-specific expression pattern of the sequence according to the invention, since no cross-hybridization was to be detected in the other types of tissue analysed (20 in total).

25

Example 7

Determination of the nucleotide sequence and derived amino acid sequence of the cDNA clone according to the invention

30 The base sequence of the longest cDNA clone according to the invention of examples 4, 5 and 6 was determined with the aid of the dideoxy method of

Sanger, F. and Coulson, A.R. "A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase" J. Mol. Biol. 94, 441 (1975). For this purpose, subclones were prepared according to example 3 and converted into single-stranded DNAs under alkaline conditions.

5

The result for the clone according to the invention is shown in SEQ ID NO: 1.

Figure 8a shows the result of the hydrophobicity plot, hydrophobic amino acid sequences being shown below the zero line. The hydrophobic regions between 10 position 620 and position 900 of the amino acid sequence which characterize the 7 transmembrane domains and lead to the conclusion of a receptor having the structure shown generally in figure 8c (from: J.D. Watson et al., "Rekombinierte DNA [Recombinant DNA]", 2nd edition, Spektrum Verlag, p. 299 (1993)) are striking. The extracellular N-terminus comprising 620 amino acids is furthermore noteworthy. A corresponding bar diagram of the protein according to the invention is shown in figure 8b. It shows the open reading frame (ORF) in the form of a box having a length of 3,114 nucleotides, which corresponds to a coding capacity of 1,038 amino acids (aa). The 7 transmembrane domains are emphasized by shading and correspond to the results 15 of the hydrophobicity analysis. The nucleotide sequence from position 3,115 to position 4,665 includes the 3' untranslated region (3' UTR) of the clone according to the invention and approximately corresponds to the cDNA clone isolated from the first cDNA library. The two cDNA clones isolated and sequenced from the two cDNA libraries of the epididymis (library 1/library 2) 20 are shown below the bar diagram. Finally, PCR subclones and the primers employed for their particular synthesis (A1, A2, E, F1) are shown underneath. The fragments used as probes are emphasized as bold lines.

Example 8

30 In situ transcript hybridization of the mRNA according to the invention in epididymal tissue sections from man and the rat

A specific probe was prepared by *in vitro* transcription using the 743 bp long fragment E - F1 (cf. ex. 6) of the DNA sequence according to the invention in the presence of [<sup>35</sup>S]CTP and was employed for determination of the 5 expression of the sequence according to the invention in frozen sections.

The results of corresponding *in situ* hybridization experiments are shown in figure 9. The white dots of the frozen sections of human origin shown in figure a using a dark-field microscope show specific signals from the *in situ* 10 hybridization in the presence of the above fragment as an "antisense" probe.

Figure b represents the result of a control experiment carried out with a comparable tissue section, in which instead of the "antisense" probe used previously, the complementary strand was used as a "sense" probe (negative control).

15 Figure a clearly shows that the DNA sequence according to the invention is expressed exclusively in the epithelial layer of the epididymis, but not in the stroma cells or muscle cells of the same human tissue.

20 Figures c and d show experiments carried out in parallel using tissue sections of the rat epididymis. Figure c shows a section of the proximal region of the epididymis, while in diagram d the distal region of the same tissue is shown. Comparing diagrams c and d, it becomes clear that the proximal region produces more intense signals, which is probably to be 25 attributed to a functional specialization of this region.

Example 9

Preparation of antigens via chemical synthesis for induction of antibodies and use of obtained antibodies in immunhistochemistry

30

Several sequence motifs within SEQ ID NO: 1 were identified by computer (DNA-

Star software package (Promega, Madison, WI, USA) as possible antigenic epitopes of the receptor protein according to the invention and synthesized chemically, together with flanking amino acids by the method of Merrifield (J.M. Stewart, "Synthesis and use of neuropeptides", Neuroendocrine Peptide

5      Methodology, ed. P.M. Conn, Academic Press, New York, p. 815-844 (1989))  
(Table I).

Table I

5 List of oligopeptides of the receptor protein used as antigens for immunization (numbers corresponding to codons of SEQ ID No 1)

Antigen	Codon No.	length of oligopeptide	localisation within the receptor protein
N1 (SEQ ID NO:3)	141-154	14 aa <sup>1)</sup>	N-terminal
N2 (SEQ ID NO:4)	266-280	15 aa	N-terminal
A (SEQ ID NO:5)	785-799	15 aa	second extracellular loop within 7TM <sup>2)</sup> domain
B (SEQ ID NO:6)	799-808	10 aa	second extracellular loop within 7TM domain
C (SEQ ID NO:7)	834-848	15 aa	third intracellular loop within 7TM domain

<sup>1)</sup> aa = amino acids

10 <sup>2)</sup> 7TM = seven transmembrane

The immunogenic polypeptides N1, N2, A, B, and C according to the invention prepared in the manner described before were each coupled to carrier 15 substances (keyhole limpet haemocyanin, KLH) using m-maleimidobenzoic acid N-hydroxysuccinamide ester (M8759, Sigma) (Sambrook et al., Molecular Cloning - A Laboratory Manual, 2nd edition, Cold Spring Harbor Laboratory Press

(1989)). Those resulting compounds were employed as antigens for immunization of rabbits, chickens and rats to generate antibodies.

Antisera obtained from rats after immunisation with A, B, and C immunogens, 5 respectively, were employed for the determination of the expression of the receptor protein according to the invention in frozen sections of human, mouse and bovine epididymis. Detection was performed by known immunohistological methods using a commercially available secondary biotinylated rabbit anti-rat antibody at a dilution of 1:100 (Dako E 468, 10 Hamburg, Germany) (Harlow and Lane, Antibodies - A Laboratory Manual, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York (1988)). The total procedure was carried out at room temperature. 10 :m cryosections were fixed in 4% paraformaldehyde for 10 min. After three washes in PBS (phosphate buffered saline pH 7.2 ) the sections were blocked in 5% normal rabbit serum in PBS. First antibody was applied for 1 h. After three washes in PBS the secondary antibody was applied for 30 min. After additional three washes in PBS, a streptavidin- biotin- horseraddish peroxidase complex (K0377, Dako, 15 Hamburg, Germany) was applied according to Sternberger et al., Immunocytochemistry, J. Wiley and Sons, New York (1979). Subsequently, the staining was carried out by incubation (Sternberger) with the chromogen diaminobenzidine (Dako, C 3467).

The results of the immunohistochemical experiments are shown in figure 10. In figures 10 a and b results of experiments are shown which were carried out 25 with tissue derived from human epididymis. The dark staining of epithelial (ep) cells of human origin shown in figure 10 a was achieved using anti-B antiserum (dilution 1:600). Figure 10 b shows the result of a control experiment carried out with a comparable tissue section, in which instead of the anti-B antiserum, the preimmune serum of the same animal was used 30 (negative control).

Figure 10 a clearly shows that the protein epitope recognized by anti-B-antiserum according to the invention is present exclusively in the epithelial layer of the epididymis, but not in stroma cells or muscle cells or within the lumen of the epididymal duct of the same human tissue.

5

Figures 10 c and d show experiments carried out in parallel using tissue sections of mouse epididymis.

Figures 10 e and f show experiments carried out in parallel using tissue 10 sections of bovine epididymis.

In figures 10 c and e the anti-B-antiserum was employed, whereas in figures 10 d and f the pre-immune serum was employed (negative controls).

15 The same results were obtained using anti-A and anti-C antisera (data not shown).

The results of this example demonstrate the following:

20 1. The use of three different antisera against different epitopes of the receptor protein, namely anti-A, anti-B or anti-C antisera, resulted in identical epithelial cell staining. This confirms the presence of the receptor protein according to the invention within the epithelium of the epididymis.

25

2. The receptor protein according to the invention is expressed in a cell type specific manner.

30 3. The high degree of homology of the receptor protein at least among mammalian species is confirmed, as the antisera against the human receptor also recognize homologues in mouse and bull.

4. The fact that antisera against A, B, and C epitopes are also reacting with mouse and bull confirm the hypothesis that the receptor protein according the invention is a "principal" structural molecule of  
5 epididymal epithelium.

The results of this example further demonstrate the extremely high cell type specificity of the receptor protein according to the invention and accordingly the suitability of the substances proposed according to the  
10 invention for specific manipulation of the functions of the epididymis in the maturation of spermatozoa. Furthermore, it can be concluded that the sequences according to the invention, including fragments thereof, can be used as specific probes for both qualitative and quantitative diagnostic investigations of biopsy samples. Moreover, the extremely high expression rate of the receptor protein according to the invention indicates its suitability as a promising mediator in the context of therapeutic or  
15 contraceptive purposes.

20 Example 10

Use of the receptor protein according to the invention and of the cDNA coding for this for isolating specific ligands.

To isolate specific ligands for the receptor protein according to the  
25 invention, the N-terminal extracellular domain (from position 1 to 620 of the amino acid sequence according to SEQ ID NO: 2) of the receptor protein was prepared in a eukaryotic expression system such as the cell lines COS-7, HEK 393 and MDCK. For this purpose, the cDNA region which codes for this domain (corresponding to position 1 to 1,860 of the nucleotide sequence according to  
30 SEQ ID NO: 1) was provided on the 3' end with a flag sequence, i.e. an oligonucleotide sequence, which codes for a known, highly specific peptide

epitope, and cloned into the expression vector pRc/CMV and pTracer-CMV (Invitrogen, San Diego, California, USA). After transfection of the cell lines with the expression vector described, the fusion product was obtained by known processes and purified by affinity chromatography using immobilized 5 antibodies directed against the flag epitope.

The fusion product was then employed as a probe in a conventional protein screening process (cf. J. Sambrook, E.F. Fritsch and T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring 10 Harbor, N.Y., chapter 2 (1989)) using a cDNA expression library from the human testis prepared in a lambda bacteriophage. For this, about 1 million independent cDNA bacterial clones which express testicular products and therefore potential ligands by IPTG induction were transferred to nylon membrane filters and incubated with the recombinant receptor-binding domain 15 under suitable conditions (see above). The receptor fragments which were not bound specifically were then removed under stringent washing conditions, so that only specific binding complexes which could be rendered visual via the flag epitope with conventional antibody detection systems (system using alkaline phosphatase, Sigma, Deisenhofen) were present on the filters. The 20 phage colonies identified in this manner were isolated and purified and subjected to a sequence determination.

In order to ascertain whether the ligands discovered can induce a signal transduction by the receptor according to the invention and are therefore 25 suitable for simulating maturation of spermatozoa in subfertile mammals, cultures of the cell lines mentioned above transfected with the total cDNA construct with and without a flag epitope on the 3' end were incubated separately with the positive ligands, and their change in the intracellular cAMP level and/or the  $Ca^{2+}$  content caused by the specific binding was 30 determined. Positive ligands which were not capable of causing such changes are suitable, for example, as antagonists for inhibition of maturation of

spermatozoa and can be employed for the preparation of contraceptive agents.

The use of such ligands for therapeutic purposes is particularly advantageous, since as a result of the tissue-specific expression of the receptor protein according to the invention, this type of tissue is influenced selectively and no side effects, or only minor side effects, are therefore to be expected.

The documents cited herein are incorporated in their entirety herein by

10 reference.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: OSTERHOFF, CAROLINE  
IVELL, RICHARD
- (ii) TITLE OF INVENTION: EPIDIDYMIS-SPECIFIC RECEPTOR PROTEIN
- (iii) NUMBER OF SEQUENCES: 11
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: NIXON & VANDERHYE P.C.
  - (B) STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
  - (C) CITY: ARLINGTON
  - (D) STATE: VA
  - (E) COUNTRY: USA
  - (F) ZIP: 22201
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: SADOFF, B.J.
  - (B) REGISTRATION NUMBER: 36,663
  - (C) REFERENCE/DOCKET NUMBER: 35-125
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 703-816-4000
  - (B) TELEFAX: 703-816-4100

(2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 4665 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..3114
- (ix) FEATURE:
  - (A) NAME/KEY: 3'UTR
  - (B) LOCATION: 3115..4665
- (ix) FEATURE:
  - (A) NAME/KEY: polyA-site
  - (B) LOCATION: 4647..4652

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

AGC CAG CCC GAG GAC GCG AGC GGC AGG TGT GCA CAG AGG TTC TCC ACT	48
Ser Gln Pro Glu Asp Ala Ser Gly Arg Cys Ala Gln Arg Phe Ser Thr	
1 5 10 15	
TTG TTT TCT GAA CTC GCG GTC AGG ATG GTT TTC TCT GTC AGG CAG TGT	96
Leu Phe Ser Glu Leu Ala Val Arg Met Val Phe Ser Val Arg Gln Cys	
20 25 30	
GGC CAT GTT GGC AGA ACT GAA GAA GTT TTA CTG ACG TTC AAG ATA TTC	144
Gly His Val Gly Arg Thr Glu Glu Val Leu Leu Thr Phe Lys Ile Phe	
35 40 45	
CTT GTC ATC ATT TGT CTT CAT GTC GTT CTG GTA ACA TCC CTG GAA GAA	192
Leu Val Ile Ile Cys Leu His Val Val Leu Val Thr Ser Leu Glu Glu	
50 55 60	
GAT ACT GAT AAT TCC AGT TTG TCA CCA CCT GCT AAA TTA TCT GTT	240
Asp Thr Asp Asn Ser Ser Leu Ser Pro Pro Ala Lys Leu Ser Val	
65 70 75 80	
GTC AGT TTT GCC CCC TCC TCC AAT GAG GTT GAA ACA ACA AGC CTC AAT	288
Val Ser Phe Ala Pro Ser Ser Asn Glu Val Glu Thr Thr Ser Leu Asn	
85 90 95	
GAT GTT ACT TTA AGC TTA CTC CCT TCA AAC GAA ACA GAA AAA ACT AAA	336
Asp Val Thr Leu Ser Leu Leu Pro Ser Asn Glu Thr Glu Lys Thr Lys	
100 105 110	
ATC ACT ATA GTA AAA ACC TTC AAT GCT TCA GGC GTC AAA CCC CAG AGA	384
Ile Thr Ile Val Lys Thr Phe Asn Ala Ser Gly Val Lys Pro Gln Arg	
115 120 125	
AAT ATC TGC AAT TTG TCA TCT ATT TGC AAT GAC TCA GCA TTT TTT AGA	432
Asn Ile Cys Asn Leu Ser Ser Ile Cys Asn Asp Ser Ala Phe Phe Arg	
130 135 140	
GGT GAG ATC ATG TTT CAA TAT GAT AAA GAA AGC ACT GTT CCC CAG AAT	480
Gly Glu Ile Met Phe Gln Tyr Asp Lys Glu Ser Thr Val Pro Gln Asn	
145 150 155 160	
CAA CAT ATA ACG AAT GGC ACC TTA ACT GGA GTC CTG TCT CTA AGT GAA	528
Gln His Ile Thr Asn Gly Thr Leu Thr Gly Val Leu Ser Leu Ser Glu	
165 170 175	
TTA AAA CGC TCA GAG CTC AAC AAA ACC CTG CAA ACC CTA AGT GAG ACT	576
Leu Lys Arg Ser Glu Leu Asn Lys Thr Leu Gln Thr Leu Ser Glu Thr	
180 185 190	
TAC TTT ATA ATG TGT GCT ACA GCA GAG GCC CAA AGC ACA TTA AAT TGT	624
Tyr Phe Ile Met Cys Ala Thr Ala Glu Ala Gln Ser Thr Leu Asn Cys	
195 200 205	
ACA TTC ACA ATA AAA CTG AAT AAT ACA ATG AAT GCA TGT GCT GCA ATA	672
Thr Phe Thr Ile Lys Leu Asn Asn Thr Met Asn Ala Cys Ala Ala Ile	
210 215 220	
GCC GCT TTG GAA AGA GTA AAG ATT CGA CCA ATG GAA CAC TGC TGC TGT	720
Ala Ala Leu Glu Arg Val Lys Ile Arg Pro Met Glu His Cys Cys Cys	

225	230	235	240	
TCT GTC AGG ATA CCC TGC CCT TCC TCC CCA GAA GAG TTG GGA AAG CTT				768
Ser Val Arg Ile Pro Cys Pro Ser Ser Pro Glu Glu Leu Gly Lys Leu				
245 250 255				
CAG TGT GAC CTG CAG GAT CCC ATT GTC TGT CTT GCT GAC CAT CCA CGT				816
Gln Cys Asp Leu Gln Asp Pro Ile Val Cys Leu Ala Asp His Pro Arg				
260 265 270				
GGC CCA CCA TTT TCT TCC AGC CAA TCC ATC CCA GTG GTG CCT CGG GCC				864
Gly Pro Pro Phe Ser Ser Gln Ser Ile Pro Val Val Pro Arg Ala				
275 280 285				
ACT GTG CTT TCC CAG GTC CCC AAA GCT ACC TCT TTT GCT GAG CCT CCA				912
Thr Val Leu Ser Gln Val Pro Lys Ala Thr Ser Phe Ala Glu Pro Pro				
290 295 300				
GAT TAT TCA CCT GTG ACC CAC AAT GTT CCC TCT CCA ATA GGG GAG ATT				960
Asp Tyr Ser Pro Val Thr His Asn Val Pro Ser Pro Ile Gly Glu Ile				
305 310 315 320				
CAA CCC CTT TCA CCC CAG CCT TCA GCT CCC ATA GCT TCC AGC CCT GCC				1008
Gln Pro Leu Ser Pro Gln Pro Ser Ala Pro Ile Ala Ser Ser Pro Ala				
325 330 335				
ATT GAC ATG CCC CCA CAG TCT GAA ACG ATC TCT TCC CCT ATG CCC CAA				1056
Ile Asp Met Pro Pro Gln Ser Glu Thr Ile Ser Ser Pro Met Pro Gln				
340 345 350				
ACC CAT GTC TCC GGC ACC CCA CCT GTG AAA GCC TCA TTT TCC TCT				1104
Thr His Val Ser Gly Thr Pro Pro Val Lys Ala Ser Phe Ser Ser				
355 360 365				
CCC ACC GTG TCT GCC CCT GCG AAT GTC AAC ACT ACC AGC GCA CCT CCT				1152
Pro Thr Val Ser Ala Pro Ala Asn Val Asn Thr Thr Ser Ala Pro Pro				
370 375 380				
GTC CAG ACA GAC ATC GTC AAC ACC AGC AGT ATT TCT GAT CTT GAG AAC				1200
Val Gln Thr Asp Ile Val Asn Thr Ser Ser Ile Ser Asp Leu Glu Asn				
385 390 395 400				
CAA GTG TTG CAG ATG GAG AAG GCT CTG TCC TTG GGC AGC CTG GAG CCT				1248
Gln Val Leu Gln Met Glu Lys Ala Leu Ser Leu Gly Ser Leu Glu Pro				
405 410 415				
AAC CTC GCA GGA GAA ATG ATC AAC CAA GTC AGC AGA CTC CTT CAT TCC				1296
Asn Leu Ala Gly Glu Met Ile Asn Gln Val Ser Arg Leu Leu His Ser				
420 425 430				
CCG CCT GAC ATG CTG GCC CCT CTG GCT CAA AGA TTG CTG AAA GTA GTG				1344
Pro Pro Asp Met Leu Ala Pro Leu Ala Gln Arg Leu Leu Lys Val Val				
435 440 445				
GAT GAC ATT GGC CTA CAG CTG AAC TTT TCA AAC ACG ACT ATA AGT CTA				1392
Asp Asp Ile Gly Leu Gln Leu Asn Phe Ser Asn Thr Thr Ile Ser Leu				
450 455 460				
ACC TCC CCT TCT TTG GCT CTG GCT GTG ATC AGA GTG AAT GCC AGT AGT				1440
Thr Ser Pro Ser Leu Ala Leu Ala Val Ile Arg Val Asn Ala Ser Ser				
465 470 475 480				

TTC AAC ACA ACT ACC TTT GTG GCC CAA GAC CCT GCA AAT CTT CAG GTT Phe Asn Thr Thr Thr Phe Val Ala Gln Asp Pro Ala Asn Leu Gln Val 485 490 495	1488
TCT CTG GAA ACC CAA GCT CCT GAG AAC AGT ATT GGC ACA ATT ACT CTT Ser Leu Glu Thr Gln Ala Pro Glu Asn Ser Ile Gly Thr Ile Thr Leu 500 505 510	1536
CCT TCA TCG CTG ATG AAT AAT TTA CCA GCT CAT GAC ATG GAG CTA GCT Pro Ser Ser Leu Met Asn Asn Leu Pro Ala His Asp Met Glu Leu Ala 515 520 525	1584
TCC AGG GTT CAG TTC AAT TTT TTT GAA ACA CCT GCT TTG TTT CAG GAT Ser Arg Val Gln Phe Asn Phe Phe Glu Thr Pro Ala Leu Phe Gln Asp 530 535 540	1632
CCT TCC CTG GAG AAC CTC TCT CTG ATC AGC TAC GTC ATA TCA TCG AGT Pro Ser Leu Glu Asn Leu Ser Leu Ile Ser Tyr Val Ile Ser Ser Ser 545 550 555 560	1680
GTT GCA AAC CTG ACC GTC AGG AAC TTG ACA AGA AAC GTG ACA GTC ACA Val Ala Asn Leu Thr Val Arg Asn Leu Thr Arg Asn Val Thr Val Thr 565 570 575	1728
TTA AAG CAC ATC AAC CCG AGC CAG GAT GAG TTA ACA GTG AGA TGT GTA Leu Lys His Ile Asn Pro Ser Gln Asp Glu Leu Thr Val Arg Cys Val 580 585 590	1776
TTT TGG GAC TTG GGC AGA AAT GGT GGC AGA GGA GGC TGG TCA GAC AAT Phe Trp Asp Leu Gly Arg Asn Gly Gly Arg Gly Gly Trp Ser Asp Asn 595 600 605	1824
GGC TGC TCT GTC AAA GAC AGG AGA TTG AAT GAA ACC ATC TGT ACC TGT Gly Cys Ser Val Lys Asp Arg Arg Leu Asn Glu Thr Ile Cys Thr Cys 610 615 620	1872
AGC CAT CTA ACA AGC TTC GGC GTT CTG CTG GAC CTA TCT AGG ACA TCT Ser His Leu Thr Ser Phe Gly Val Leu Leu Asp Leu Ser Arg Thr Ser 625 630 635 640	1920
GTG CTG CCT GCT CAA ATG ATG GCT CTG ACG TTC ATT ACA TAT ATT GGT Val Leu Pro Ala Gln Met Met Ala Leu Thr Phe Ile Thr Tyr Ile Gly 645 650 655	1968
TGT GGG CTT TCA TCA ATT TTT CTG TCA GTG ACT CTT GTA ACC TAC ATA Cys Gly Leu Ser Ser Ile Phe Leu Ser Val Thr Leu Val Thr Tyr Ile 660 665 670	2016
GCT TTT GAA AAG ATC CGG AGG GAT TAC CCT TCC AAA ATC CTC ATC CAG Ala Phe Glu Lys Ile Arg Arg Asp Tyr Pro Ser Lys Ile Leu Ile Gln 675 680 685	2064
CTG TGT GCT CTG CTT CTG CTG AAC CTG GTC TTC CTC CTG GAC TCG Leu Cys Ala Ala Leu Leu Leu Asn Leu Val Phe Leu Leu Asp Ser 690 695 700	2112
TGG ATT GCT CTG TAT AAG ATG CAA GGC CTC TGC ATC TCA GTG GCT GTA Trp Ile Ala Leu Tyr Lys Met Gln Gly Leu Cys Ile Ser Val Ala Val 705 710 715 720	2160
TTT CTT CAT TAT TTT CTC TTG GTC TCA TTC ACA TGG ATG GGC CTA GAA Phe Leu His Tyr Phe Leu Leu Val Ser Phe Thr Trp Met Gly Leu Glu	2208

725	730	735	
GCA TTC CAT ATG TAC CTG GCC CTT GTC AAA GTA TTT AAT ACT ACT TAC ATC Ala Phe His Met Tyr Leu Ala Leu Val Lys Val Phe Asn Thr Tyr Ile 740	745	750	2256
CGA AAA TAC ATC CTT AAA TTC TGC ATT GTC GGT TGG GGG GTA CCA GCT Arg Lys Tyr Ile Leu Lys Phe Cys Ile Val Gly Trp Gly Val Pro Ala 755	760	765	2304
GTG GTT GTG ACC ATC ATC CTG ACT ATA TCC CCA GAT AAC TAT GGG CTT Val Val Val Thr Ile Ile Leu Thr Ile Ser Pro Asp Asn Tyr Gly Leu 770	775	780	2352
GGA TCC TAT GGG AAA TTC CCC AAT GGT TCA CCG GAT GAC TTC TGC TGG Gly Ser Tyr Gly Lys Phe Pro Asn Gly Ser Pro Asp Asp Phe Cys Trp 785	790	795	2400
ATC AAC AAC AAT GCA GTA TTC TAC ATT ACG GTG GTG GGA TAT TTC TGT Ile Asn Asn Asn Ala Val Phe Tyr Ile Thr Val Val Gly Tyr Phe Cys 805	810	815	2448
GTG ATA TTT TTG CTG AAC GTC AGC ATG TTC ATT GTG GTC CTG GTT CAG Val Ile Phe Leu Leu Asn Val Ser Met Phe Ile Val Val Leu Val Gln 820	825	830	2496
CTC TGT CGA ATT AAA AAG AAG AAG CAA CTG GGA GCC CAG CGA AAA ACC Leu Cys Arg Ile Lys Lys Lys Gln Leu Gly Ala Gln Arg Lys Thr 835	840	845	2544
AGT ATT CAA GAC CTC AGG AGT ATC GCT GGC CTT ACA TTT TTA CTG GGA Ser Ile Gln Asp Leu Arg Ser Ile Ala Gly Leu Thr Phe Leu Leu Gly 850	855	860	2592
ATA ACT TGG GGC TTT GCC TTC TTT GCC TGG GGA CCA GTT AAC GTG ACC Ile Thr Trp Gly Phe Ala Phe Ala Trp Gly Pro Val Asn Val Thr 865	870	875	2640
TTC ATG TAT CTG TTT GCC ATC TTT AAT ACC TTA CAA GGA TTT TTC ATA Phe Met Tyr Leu Phe Ala Ile Phe Asn Thr Leu Gln Gly Phe Phe Ile 885	890	895	2688
TTC ATC TTT TAC TGT GTG GCC AAA GAA AAT GTC AGG AAG CAA TGG AGG Phe Ile Phe Tyr Cys Val Ala Lys Glu Asn Val Arg Lys Gln Trp Arg 900	905	910	2736
CGG TAT CTT TGT TGT GGA AAG TTA CGG CTG GCT GAA AAT TCT GAC TGG Arg Tyr Leu Cys Cys Gly Lys Leu Arg Leu Ala Glu Asn Ser Asp Trp 915	920	925	2784
AGT AAA ACT GCT ACT AAT GGT TTA AAG AAG CAG ACT GTA AAC CAA GGA Ser Lys Thr Ala Thr Asn Gly Leu Lys Lys Gln Thr Val Asn Gln Gly 930	935	940	2832
GTC TCC AGC TCT TCA AAT TCC TTA CAG TCA AGC AGT AAC TCC ACT AAC Val Ser Ser Ser Asn Ser Leu Gln Ser Ser Asn Ser Thr Asn 945	950	955	2880
TCC ACC ACA CTG CTA GTG AAT AAT GAT TGC TCA GTA CAC GCA AGC GGG Ser Thr Thr Leu Leu Val Asn Asn Asp Cys Ser Val His Ala Ser Gly 965	970	975	2928

AAT GGA AAT GCT TCT ACA GAG AGG AAT GGG GTC TCT TTT AGT GTT CAG Asn Gly Asn Ala Ser Thr Glu Arg Asn Gly Val Ser Phe Ser Val Gln 980 985 990	2976
AAT GGA GAT GTG TGC CTT CAC GAT TTC ACT GGA AAA CAG CAC ATG TTT Asn Gly Asp Val Cys Leu His Asp Phe Thr Gly Lys Gln His Met Phe 995 1000 1005	3024
AAC GAG AAG GAA GAT TCC TGC AAT GGG AAA GGC CGT ATG GCT CTC AGA Asn Glu Lys Glu Asp Ser Cys Asn Gly Lys Gly Arg Met Ala Leu Arg 1010 1015 1020	3072
AGG ACT TCA AAG CGG GGA AGC TTA CAC TTT ATT GAG CAA ATG Arg Thr Ser Lys Arg Gly Ser Leu His Phe Ile Glu Gln Met 1025 1030 1035	3114
TGATTCCCTT CTTCTAAAAT CAAAGCATGA TGCTTGACAG TGTGAAATGT CCAATTTCAC CTTTTACACA ATGTGAGATG TATGAAAATC AACTCATTTC ATTCTCGGCA ACATCTGGAG	3174
AAGCATAAGC TAATTAAGGG CGATGATTAT TATTACAAGA AGAAACCAAG ACATTACACC ATGGTTTTA GACATTCTG ATTTGGTTTC TTATCTTCA TTTTATAAGA AGGTTGGTTT	3234
TAAACAATAC ACTAAGAATG ACTCCTATAA AGAAAACAAA AAAAGGTAGT GAACTTTCAG CTACCTTTA AAGAGGCTAA GTTATCTTG ATAACATCAT ATAAAGCAAC TGTTGACTTC	3294
AGCCTGTTGG TGAGTTAGT TGTGCATGCC TTTGTTGTAT ATAAGCTAAA TTCTAGTGAC CCATGTGTCA AAAATCTTAC TTCTACATTT TTTGTTATTT ATTTTCTACT GTGTAATGT	3354
ATTCCCTTGT AGAACATGG TTGTTTGTC TCACGTGATA ATTCAAGAAAA TCCTTGCTCG TTCCGCAAAT CCTAAAGCTC CTTTGGAGA TGATATAGGA TGTGAAATAC AGAAACCTCA	3414
GTGAAATCAA GAAATAATGA TCCCAGCCAG ACTGAGAAAA TGTAAGCAGA CAGTGCACACA GTTAGCTCAT ACAGTGCCTT TGAGCAAGTT AGGAAAAGAT GCCCCCACGT GGCAGACACA	3474
GCCCTATGGG TCATGGTTG ACAAACAGAG TGAGAGACCA TATTTAGCC CCACTCACCC TCTTNGGTGC ACGACCTGTA CAGCCAAACA CAGCATCCAA TATGAATACC CATCCCCTGA	3534
CCGCATCCCC AGTAGTCAGA TTATAGAAC TGCACCAAGA TGTGAAATAC AGAAACCTCA AGGTACATAC ACTGCCGCTT CTCAAATCCC CAGAGCCTT AGGAACAGGA GAGTAGACTA	3594
GGATTCCCTTC TCTTAAAAAG GTACATATAT ATGGAAAAAA ATCATATTGC CGTTCTTAA AAGGCAACTG CATGGTACAT TGTTGATTGT TATGACTGGT ACACTCTGGC CCAGCCAGAG	3654
CTATAATTGT TTTTAAATG TGTCTTGAAG AATGCCACAGT GANAAGGGGA GTAGCTATTG GGAACAGGGGA ACTGCTTAC ACTGCTATTG TTGCTACATG TATCGAGCCT TGATTGCTCC	4014
TAGTTATATA CAGGGTCTAT CTTGCTTCCT ACCTACATCT GCTTGAGCAG TGCCTCAAGT ACATCCTTAT TAGAACATT TCAAACCCCT TTTAGTTAAG TCTTCACTA AGGTTCTCTT	4074
	4134
	4194
	4254
	4314
	4374
	4434
	4494

GCATATATTT CAAGTGAATG TTGGATCTCA GACTAACCAT AGTAATAATA CACATTCTG	4554
TGAGTGCTGA CTTGTCTTG CAATATTCT TTTCTGATTT ATTTAATTCTT CTTGTATTAA	4614
TATGTTAAAA TCAAAAATGT TAAAATCAAT GAAATAAATT TGCAGTTAAG A	4665

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1038 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Ser Gln Pro Glu Asp Ala Ser Gly Arg Cys Ala Gln Arg Phe Ser Thr	
1 5 10 15	
Leu Phe Ser Glu Leu Ala Val Arg Met Val Phe Ser Val Arg Gln Cys	
20 25 30	
Gly His Val Gly Arg Thr Glu Glu Val Leu Leu Thr Phe Lys Ile Phe	
35 40 45	
Leu Val Ile Ile Cys Leu His Val Val Leu Val Thr Ser Leu Glu Glu	
50 55 60	
Asp Thr Asp Asn Ser Ser Leu Ser Pro Pro Ala Lys Leu Ser Val	
65 70 75 80	
Val Ser Phe Ala Pro Ser Ser Asn Glu Val Glu Thr Thr Ser Leu Asn	
85 90 95	
Asp Val Thr Leu Ser Leu Leu Pro Ser Asn Glu Thr Glu Lys Thr Lys	
100 105 110	
Ile Thr Ile Val Lys Thr Phe Asn Ala Ser Gly Val Lys Pro Gln Arg	
115 120 125	
Asn Ile Cys Asn Leu Ser Ser Ile Cys Asn Asp Ser Ala Phe Phe Arg	
130 135 140	
Gly Glu Ile Met Phe Gln Tyr Asp Lys Glu Ser Thr Val Pro Gln Asn	
145 150 155 160	
Gln His Ile Thr Asn Gly Thr Leu Thr Gly Val Leu Ser Leu Ser Glu	
165 170 175	
Leu Lys Arg Ser Glu Leu Asn Lys Thr Leu Gln Thr Leu Ser Glu Thr	
180 185 190	
Tyr Phe Ile Met Cys Ala Thr Ala Glu Ala Gln Ser Thr Leu Asn Cys	
195 200 205	
Thr Phe Thr Ile Lys Leu Asn Asn Thr Met Asn Ala Cys Ala Ala Ile	
210 215 220	

Ala Ala Leu Glu Arg Val Lys Ile Arg Pro Met Glu His Cys Cys Cys  
225 230 235 240

Ser Val Arg Ile Pro Cys Pro Ser Ser Pro Glu Glu Leu Gly Lys Leu  
245 250 255

Gln Cys Asp Leu Gln Asp Pro Ile Val Cys Leu Ala Asp His Pro Arg  
260 265 270

Gly Pro Pro Phe Ser Ser Gln Ser Ile Pro Val Val Pro Arg Ala  
275 280 285

Thr Val Leu Ser Gln Val Pro Lys Ala Thr Ser Phe Ala Glu Pro Pro  
290 295 300

Asp Tyr Ser Pro Val Thr His Asn Val Pro Ser Pro Ile Gly Glu Ile  
305 310 315 320

Gln Pro Leu Ser Pro Gln Pro Ser Ala Pro Ile Ala Ser Ser Pro Ala  
325 330 335

Ile Asp Met Pro Pro Gln Ser Glu Thr Ile Ser Ser Pro Met Pro Gln  
340 345 350

Thr His Val Ser Gly Thr Pro Pro Val Lys Ala Ser Phe Ser Ser  
355 360 365

Pro Thr Val Ser Ala Pro Ala Asn Val Asn Thr Thr Ser Ala Pro Pro  
370 375 380

Val Gln Thr Asp Ile Val Asn Thr Ser Ser Ile Ser Asp Leu Glu Asn  
385 390 395 400

Gln Val Leu Gln Met Glu Lys Ala Leu Ser Leu Gly Ser Leu Glu Pro  
405 410 415

Asn Leu Ala Gly Glu Met Ile Asn Gln Val Ser Arg Leu Leu His Ser  
420 425 430

Pro Pro Asp Met Leu Ala Pro Leu Ala Gln Arg Leu Leu Lys Val Val  
435 440 445

Asp Asp Ile Gly Leu Gln Leu Asn Phe Ser Asn Thr Thr Ile Ser Leu  
450 455 460

Thr Ser Pro Ser Leu Ala Leu Ala Val Ile Arg Val Asn Ala Ser Ser  
465 470 475 480

Phe Asn Thr Thr Phe Val Ala Gln Asp Pro Ala Asn Leu Gln Val  
485 490 495

Ser Leu Glu Thr Gln Ala Pro Glu Asn Ser Ile Gly Thr Ile Thr Leu  
500 505 510

Pro Ser Ser Leu Met Asn Asn Leu Pro Ala His Asp Met Glu Leu Ala  
515 520 525

Ser Arg Val Gln Phe Asn Phe Phe Glu Thr Pro Ala Leu Phe Gln Asp  
530 535 540

Pro Ser Leu Glu Asn Leu Ser Leu Ile Ser Tyr Val Ile Ser Ser Ser  
545 550 555 560

Val Ala Asn Leu Thr Val Arg Asn Leu Thr Arg Asn Val Thr Val Thr  
565 570 575

Leu Lys His Ile Asn Pro Ser Gln Asp Glu Leu Thr Val Arg Cys Val  
580 585 590

Phe Trp Asp Leu Gly Arg Asn Gly Gly Arg Gly Trp Ser Asp Asn  
595 600 605

Gly Cys Ser Val Lys Asp Arg Arg Leu Asn Glu Thr Ile Cys Thr Cys  
610 615 620

Ser His Leu Thr Ser Phe Gly Val Leu Leu Asp Leu Ser Arg Thr Ser  
625 630 635 640

Val Leu Pro Ala Gln Met Met Ala Leu Thr Phe Ile Thr Tyr Ile Gly  
645 650 655

Cys Gly Leu Ser Ser Ile Phe Leu Ser Val Thr Leu Val Thr Tyr Ile  
660 665 670

Ala Phe Glu Lys Ile Arg Arg Asp Tyr Pro Ser Lys Ile Leu Ile Gln  
675 680 685

Leu Cys Ala Ala Leu Leu Leu Asn Leu Val Phe Leu Leu Asp Ser  
690 695 700

Trp Ile Ala Leu Tyr Lys Met Gln Gly Leu Cys Ile Ser Val Ala Val  
705 710 715 720

Phe Leu His Tyr Phe Leu Leu Val Ser Phe Thr Trp Met Gly Leu Glu  
725 730 735

Ala Phe His Met Tyr Leu Ala Leu Val Lys Val Phe Asn Thr Tyr Ile  
740 745 750

Arg Lys Tyr Ile Leu Lys Phe Cys Ile Val Gly Trp Gly Val Pro Ala  
755 760 765

Val Val Val Thr Ile Ile Leu Thr Ile Ser Pro Asp Asn Tyr Gly Leu  
770 775 780

Gly Ser Tyr Gly Lys Phe Pro Asn Gly Ser Pro Asp Asp Phe Cys Trp  
785 790 795 800

Ile Asn Asn Asn Ala Val Phe Tyr Ile Thr Val Val Gly Tyr Phe Cys  
805 810 815

Val Ile Phe Leu Leu Asn Val Ser Met Phe Ile Val Val Leu Val Gln  
820 825 830

Leu Cys Arg Ile Lys Lys Lys Gln Leu Gly Ala Gln Arg Lys Thr  
835 840 845

Ser Ile Gln Asp Leu Arg Ser Ile Ala Gly Leu Thr Phe Leu Leu Gly  
850 855 860

Ile Thr Trp Gly Phe Ala Phe Phe Ala Trp Gly Pro Val Asn Val Thr  
865 870 875 880

Phe Met Tyr Leu Phe Ala Ile Phe Asn Thr Leu Gln Gly Phe Phe Ile

	885	890	895
Phe Ile Phe Tyr Cys Val Ala Lys Glu Asn Val Arg Lys Gln Trp Arg			
900	905	910	
Arg Tyr Leu Cys Cys Gly Lys Leu Arg Leu Ala Glu Asn Ser Asp Trp			
915	920	925	
Ser Lys Thr Ala Thr Asn Gly Leu Lys Lys Gln Thr Val Asn Gln Gly			
930	935	940	
Val Ser Ser Ser Ser Asn Ser Leu Gln Ser Ser Asn Ser Thr Asn			
945	950	955	960
Ser Thr Thr Leu Leu Val Asn Asn Asp Cys Ser Val His Ala Ser Gly			
965	970	975	
Asn Gly Asn Ala Ser Thr Glu Arg Asn Gly Val Ser Phe Ser Val Gln			
980	985	990	
Asn Gly Asp Val Cys Leu His Asp Phe Thr Gly Lys Gln His Met Phe			
995	1000	1005	
Asn Glu Lys Glu Asp Ser Cys Asn Gly Lys Gly Arg Met Ala Leu Arg			
1010	1015	1020	
Arg Thr Ser Lys Arg Gly Ser Leu His Phe Ile Glu Gln Met			
1025	1030	1035	

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 14 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Ala Phe Phe Arg Gly Glu Ile Met Phe Gln Tyr Asp Lys Glu  
1 5 10

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Cys Leu Ala Asp His Pro Arg Gly Pro Pro Phe Ser Ser Ser Gln  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Gly Ser Tyr Gly Lys Phe Pro Asn Gly Ser Pro Asp Asp Phe Cys  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Cys Trp Ile Asn Asn Asn Ala Val Phe Tyr  
1 5 10

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Cys Arg Ile Lys Lys Lys Gln Leu Gly Ala Gln Arg Lys Thr  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGCTATGGGA GCTGAAG

17

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGTCAATGGC AGGGCTG

17

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CATCCGAAAAA TACATCC

17

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGAAGGCACA CATCTCC

17